Summary of the addition reading

Paper Title: Identifying the Activities Supported by Locations with Community-Authored Content

1. Summary

Nowadays, there are a lot of location-based applications with community-authored content, such as location specific reviews, which could offer a lot of useful information. So this study focuses on the Yelp’s community-authored reviews to identify a set of potential activities.

Context-aware applications commonly require knowledge of a person’s location and activity based on relying on low cost sensors meanwhile this paper proposes a new way to identify potential activities by the locations a person may visit by processing community-authored reviews on Yelp. They take use of such information to build a corpus of potential activities for each location. Although, the activities they identified might not be related to that person immediately, it shows the potential possibility that that person would do in that location. In this study, they also do the evaluations of precision by user study, which shows following points:

1. Yelp reviews can be processed to identify the activities supported by the reviewed location with nice precision up to 79.3% and recall up to 55.9%.

2. The number of reviews authored for a location has a significant impact on precision.

3. There might be a difference in the activities which user identified with the reviews identified for a location.

Yelp is an online community whose members author reviews about specific public and commercial locations that could provide rich information about a location by the community-authored content. And Yelp’s community-authored content often has a specific location tag and activity tag like “take rowboat”. So the first step of study is to deriving activities from community-authored content by four steps:

1. Harvest the review texts and related attributes.

2. Parse the review texts to identify each sentence.

3. Tag each word of a sentence with its part of speech and extract local verb-noun pairs to form activities.

4. Populate and update the activity database with the identified verb-noun pairs.

The next step is to do the experimental validation of the verb-noun pair dataset. So the participants were required to provide activities that they have done at one of the locations and validate the verb-noun pairs as actual activities supported by the respective location, which could get a list of those common verb-noun pairs with results of true positive, false negative, and false positive. In this study, we evaluated three methods of filtering the verb-noun pairs including: no filter, 1st person, and ƒ>1. And finally this study generate a result with some figures which shows very good precision.

In short, Yelp’s community-authored reviews offer a rich data source that can be processed to identify activities by generating the verb-noun pairs that are supported by a location with great precision and the result of this study could be used on a breadth of context-aware applications.

2. Discussion

Certainly, there are some cons and pros of this study. One of the pros is that this study is gains good mean precision of up to 79.3% and recall up to 55.9% across the 14 locations. And one of the cons is that the frequency analysis is too simple, which causes too much false positives in result, and this study might need some more sophisticated methods of natural language processing to deal with this problem. Also another of the cons might be the small sample size. Because the number of reviews authored for a location has a strong influence on precision, which encourages the study using a bigger sample size for each location in the future. The last point of the cons is that this study didn’t take the timeliness of an activity into consideration, but the timeliness of an activity has influence on activities validity.

Furthermore, the context-aware services that provide more relevant services to the user based on current location of a mobile user could take use of the result of this study. Not only two examples provided in the paper, there are more samples flood into my head about the context-aware services. For example, we could use the result to do some prediction of where the user to go and do the recommendation. Also this prediction could combine more information like the time of the day, temperature and so on. Another great area of using the context-aware services is health care. Based on the community-authored information and hospital staff’s comment or feedback, which could speed up the process of being in treatment meantime the patients might be better taken care of through these information.